

VARIATION DANS LES ACCORDS DU FRANÇAIS INCLUSIF*

Yarubi Díaz and David Heap
Western University

1. Introduction

Si bien la tradition grammaticale française a longtemps maintenu le contraire, un ensemble croissant de recherches psycholinguistiques sur le français confirment qu'il est très difficile de ne pas associer les marques d'accord au masculin avec des sujets animés de genre (social) masculin (Brauer et Landry 2008). La poursuite d'une représentation linguistique plus équitable des femmes, et plus récemment de la diversité des genres non-binaires, implique un défi particulier pour les langues à genre grammatical marqué, comme le français (Dister et Moreau 2009, Viennot et al. 2015). Ainsi, le langage inclusif (désormais LI) propose depuis des décennies une variété d'alternatives graphiques au masculin générique traditionnel telles que le marquage du féminin avec du soulignement (étudiantes) ou des majuscules (étudiantEs); l'utilisation des doublets complets, en ordre masculin-féminin (étudiants et étudiantes) ou féminin-masculin (étudiantes et étudiants) et aussi des doublets abrégés à l'emploi de différentes formes typographiques tels que le tiret (étudiant-es), le point (étudiant.es), le point médian (étudiant·es), les barres obliques (étudiant/e), entre autres procédés, tels que l'emploi de formes épïcènes ou de formes alternatives (néologismes).

En même temps, ces choix ont des répercussions directes dans le système de concordance du français qui semble combiner ces différentes alternatives, particulièrement en contexte non institutionnel. Cette étude vise à déterminer les patrons caractéristiques des accords en français inclusif non-institutionnel et à comparer des emplois dans différentes régions francophones. Pour ceci, nous analysons un corpus de Twitter à partir d'une méthodologie quantitative ciblée sur la fréquence de combinaisons et les localisations d'origine.

* Nous sommes reconnaissant.e.s des commentaires constructifs du public à l'ACL 2020 ainsi que de Heather Burnett, Céline Pozniac, Jeff Tennant et Naomi Nagy. Cette recherche a bénéficié du soutien de subventions Social Sciences and Humanities Research Council Explore Grant et Faculty Research Development Fund de l'Université Western Ontario.

1.1 Twitter : français non institutionnel et études de variation

Notre intérêt est d'étudier un français moins attaché à la formalité des institutions et qui soit, pour ainsi dire, un français plus malléable aux intérêts des personnes qui l'utilisent, comme celui de réseaux sociaux. Alpheratz affirme que « le français inclusif est un phénomène qui est d'abord lié aux situations d'énonciation propres au XXI^e siècle, lequel se distingue par cette (r)évolution technologique majeure que représente le passage de l'écrit au numérique » (2018: 4). Dans ce sens, nous proposons une analyse du LI sur une des plateformes à plus ample usage, Twitter. Trois caractéristiques le positionnent comme un contexte idéal pour observer les accords du LI et leur variation : la spontanéité, la brièveté et le libre accès. Du fait que la pression normative institutionnelle est basse, les tweets présentent un niveau de spontanéité élevé puisqu'ils demandent moins de temps et d'effort que d'autres textes plus traditionnels. La brièveté des messages, limités à 280 caractères, tend à favoriser la variation à partir de toutes sortes de raccourcis. Le libre accès de Twitter aux internautes permet en même temps la prise de données pour l'étude du LI dans le monde francophone.

Des études de variation comme celles de Bouzouita et Pato (2019), Marttinen Larsson et Bouzouita (2018) et Escalante (2015) constatent la pertinence d'emploi de Twitter pour mener des études de variation: un corpus de ce type présente l'avantage de permettre d'analyser des données écrites non-institutionnelles souvent difficiles à recueillir ailleurs.

1.2 Études antérieures sur le LI en français

Abbou (2017) explore les formes graphiques du LI prenant en compte l'axe idéologique. Elle analyse des pamphlets anarchistes en combinaison avec des entretiens et ses résultats établissent des associations idéologiques avec certains des procédés typographiques inclusifs. La majuscule était vue par les personnes interviewées comme un procédé dépassé, alors que le genre féminin marqué par les parenthèses était « inacceptable politiquement », étant perçu comme preuve d'attachement au gouvernement de droite. Les autres procédés typographiques n'ont pas particulièrement montré d'association idéologique (2017: 65).

Burnett et Pozniak (2020) étudient les associations entre les différents procédés de LI dans 871 brochures de licence de 12 universités parisiennes, d'après les disciplines, le prestige (positionnement idéologique) et le taux de parité de genre de chaque institution. Des sept procédés de LI attestés, les auteures ont trouvé qu'il y a quand même quatre qui démontrent une association statistiquement significative: le point médian est associé avec les universités perçues comme moins prestigieuses ou plus militantes, et c'est également la forme la plus fréquente. Les parenthèses s'associent avec les universités les plus prestigieuses qui sont également celles qui sont les moins militantes. De sa part, les tirets sont évités dans les départements avec plus de femmes. Le point est la forme non marquée utilisée à travers tous les domaines et c'est également le forme qui apparaît dans tous les domaines (2020: 26).

Pour ce qui est spécifiquement des accords dans le LI, Alpheratz (2018) distingue trois types : l'accord de proximité, lequel jusqu'à date relativement récente était permis en

français comme dans les langues classiques: on peut accorder avec le substantif le plus proche: « les hommes et les femmes sont belles » (2018: 247). Il y a aussi un accord par majorité: il consiste à faire coïncider le genre grammatical avec le genre des référents majoritaires: « trois femmes et un homme sont les grandes gagnantes de cette édition » (2018: 248). Pour sa part, l'accord d'énonciation, comme dans la phrase « **elle y a** de quoi vivre longtemps et sans manquer de rien » (2018: 254) consiste à accorder le sujet d'une structure impersonnelle avec le genre perçu ou supposé de l'énonciataire. Alpheratz remarque qu'il n'y pas pour le moment de registres où ce type d'accord est utilisé en dehors de textes de fiction littéraire (2018: 254).

Du fait que notre étude vise à déterminer les patrons et la variation des accords inclusifs, nos analyses ne comportent pas, à ce niveau, d'entretiens ou d'autres sources pour avoir accès aux processus de choix des énonciateurs et à leurs motivations. De cette manière, il n'y a pas d'éléments qui nous permettent de savoir si un accord de l'adjectif au masculin pluriel est motivé par un accord de proximité, de majorité ou même d'énonciation. Alors, la classification d'Alpheratz ne semble pas être la plus adéquate pour notre corpus. Les patrons trouvés seront plutôt classifiés en accords uniformes et accords mixtes (voir 2.3).

2. Méthodologie

Le corpus n'est pas représentatif du français sur Twitter en général dans le sens qu'il contient uniquement des tweets à LI, laissant de côté la grande masse des emplois non inclusifs, ce qui se justifie par notre objectif d'étudier les formes d'accords inclusifs. Étant donné les limitations de cette étude, nous n'avons pas abordé de dimensions idéologiques ni thématiques dans ce travail. Un des buts de cette étude est de contribuer à la construction de méthodologies pour l'étude du LI en contexte non institutionnel, raison pour laquelle nous avons construit un corpus de Twitter à partir du logiciel Netlytic.org. Ce logiciel recueille automatiquement de gros volumes de texte sur de sites de médias sociaux à partir d'un ensemble de paramètres. Dans le cas de la prise de notre échantillon, il a été nécessaire d'établir : la langue de tweets, les pays francophones et les descripteurs ou termes de recherche. Par la suite, nous présenterons des détails concernant ces trois derniers.

2.1 Les pays francophones

Nous avons pris comme base une liste hiérarchisant les pays à haut pourcentage de francophones proposée par l'*Observatoire de la langue française de l'Organisation internationale de la Francophonie* (2018) et nous l'avons combinée avec une classification des pays selon leur usage d'internet et leur accès aux réseaux sociaux (InternetWorldStats 2020). Puisque le paramètre de recherche de Netlytic ne fonctionne qu'avec des villes, nous nous avons tenté de prendre de données surtout dans les capitales ou les grandes villes des pays francophones sélectionnés : Alger, Abidjan, Bamako, Bruxelles, Casablanca, Cayenne, Dakar, Genève, Hanoi, Kigali, Kinshasa, Luxembourg, Maroc, Monaco, Montréal, Ottawa, Paris, Port-au-Prince, Québec, Yaoundé et Yamoussoukro. À partir de résultats de ce recueil, nous avons choisi les huit pays présentés dans le Tableau 1 en ordre

alphabétique, accompagnés de villes et de rayons de recherche utilisés. Le choix de ces pays présentés se doit à la quantité et à la pertinence de données pour la composition de l'échantillon. En d'autres mots, nous avons utilisé pour cette étude les pays dont les villes ont généré le plus de résultats à partir de mécanismes de recherche utilisés.

Tableau 1. Les 8 pays francophones sélectionnés et les rayons de recherche utilisés

Pays	Ville(s)	Rayon de recherche
Algérie	Alger, Bejaïa, Kabylie	200km
Belgique	Bruxelles	100km
Cameroun	Youandé, Douala	100km
Canada	Québec, Montréal, Ottawa	100km
France	Paris	100km
Luxembourg	Ville de Luxembourg	10km
République Dém du Congo	Kinshasa	100km
Suisse	Genève	10km

Pour diminuer les erreurs de localisation, nous avons testé de rayons aux kilométrages différents pour notre recherche et nous avons fini par choisir le rayon qui était le plus restreint mais qui, en même temps, donnait le plus de résultats favorables (voir tableau 1). Ainsi, les rayons de recherche de villes plus petites sont plus petits que ceux de villes à plus grande extension.

2.2 Les descripteurs

Nous avons élaboré un ensemble constitué de 105 descripteurs divisés en trois catégories : des doublets complets, des doublets abrégés et des formes alternatives. Dans le premier groupe, nous avons utilisé des articles indéfinis et définis; des pronoms personnels, des substantifs et des pronoms et adjectifs indéfinis. Dans le deuxième groupe, nous avons utilisé des formes abrégées des déterminants indéfinis; des pronoms et des adjectifs indéfinis et des professions et métiers. Pour les formes alternatives, nous avons choisi quelques pronoms, noms et adjectifs qui constituent de néologismes. Le tableau 2 présente la liste complète de descripteurs utilisés pour nos recherches sur Twitter. Cette sélection comporte des mots du LI recueillis de Haddad et Baric (2017), Lavieenqueer (2018) et Alpheratz (2018). Avec cette sélection, nous avons voulu faire un tour d'horizon initial pour trouver les formes les plus susceptibles de donner des résultats à travers la francophonie.

Tableau 2. Les 105 descripteurs utilisés pour la prise de l'échantillon

1	2	3	4	5	6	7
Doublets abrégés	Doublets complets	Doublets complets	Doublets complets	Doublets complets	Formes alternatives	Doublets abrégés
Pronoms, adjectifs et articles indéfinis	Pronoms et adjectifs indéfinis	Déterminants: définis, indéfinis, possessifs	Pronoms sujets	Substantifs	Pronoms, noms, adjectifs néoformes	Métiers et professions
certain·e	certain et	un et une	il et elle;	les hommes et	iel	chargé·e
chacun·e	certaine	une et un	elle et lui;	les femmes ; les	iels	député·e
un·e	certaine et	le et la	ils et elles ;	femmes et les	celleux	élu·e
tout·e	certain	la et le	elles et eux ;	hommes; les	adelphe	médical·e
certain·e	chacune et	ton et ta	ceux et celles;	filles et les	epoux	syndical·e
chacun·e	chacun	ta et ton	celles et ceux;	garçons; les	partenaire	chef·fe
un·e	certaines et	un ou une	il ou elle; elle	garçons et les	heureux	professionnel·le
tout·e	certain	une ou un	ou lui;	filles; les	jaloux	citoyen·ne
certain·es	certaines et	le ou la	ils ou elles;	hommes ou les	joueuse	lycéen·ne
chacun·es	certaines	la ou le	elles ou eux;	femmes; les	doux	technicien·ne
tou·tes	toutes et tous	ton ou ta	ceux ou celles;	femmes ou les	acteurice	chercheur·e
certain·e·s	tous et toutes	ta ou ton	celles ou ceux;	hommes;	autaire	entrepreneur·e
chacun·e·s	certain ou	son ou sa	celui ou celle;	les filles ou les	danseuse	ingénieur·e
tou·t·e·s	certaine	sa ou son	celle ou celui	garçons;	chanteuse	professeur·e
certain·e·s	certaine ou			les garçons ou	belleaux	professeuse
chacun·e·s	certain			les filles	belleau	adjoint·e
tou·te·s	chacune ou					agent·e
	chacun					assistant·e
	certaines ou					avocat·e
	certain					consultant·e
	certaines ou					étudiant·e
	certaines					
	toutes ou tous					
	tous ou toutes					

2.3 L'échantillon et procédures de dépouillage

Nous avons recueilli un échantillon de 13.613 tweets dont la distribution par pays est présentée dans la première ligne du tableau 3. Toutes les occurrences ont été prises le 12 mai 2020 et en suivant les mêmes paramètres de recherche. Le dépouillement inclut l'élimination de doublons, des retweets, des tweets sans éléments explicites de localisation et ceux avec des localisations irréelles/invalides comme « Heaven », « dans les étoiles », « Narnia », etc. Nous avons aussi supprimé les tweets sans pertinence, c'est-à-dire, ceux qui n'avaient pas d'éléments de LI. C'est à partir de ce travail que nous sommes arrivés à avoir un corpus définitif de 1.504 tweets. Le tableau 3 montre l'échantillon, le processus de dépouillement par pays et le corpus définitif:

Tableau 3. Échantillon, dépouillement et corpus

Pays francophone	ALG	BRU	CAM	CAN	FRA	LUX	RDC	SUI	Total
Tweets extraits par Netlytic: échantillon	67	916	2000	1076	6218	51	3133	152	13613
Doublons supprimés	38	569	1367	599	4738	28	2030	59	9338
Tweets à localisation absente	0	22	228	7	113	0	406	2	778
Tweets avec erreur de localisation	1	22	394	64	60	1	643	5	1190
Tweets sans pertinence (pas de LI)	8	61	6	93	471	9	32	33	713
Corpus	20	242	5	313	836	13	22	53	1504

Le processus de dépouillement a été complexe, ce qui se voit dans la large différence entre l'échantillon initial et le corpus. Cette réduction a particulièrement affecté les données issues du Cameroun et du Luxembourg dont le corpus n'a pas suffisamment d'occurrences pour pouvoir faire des analyses significatives.

2.4 Corpus: codage et classification des accords

Un système de codage a été préparé pour typifier les procédés d'accord que nous avons trouvés dans les tweets composant notre corpus, ce que détaillons dans le tableau 4:

Tableau 4. Codage des procédés d'accord d'inclusion dans le corpus

M	accord au masculin	T	accord avec trait-d'union
F	accord au féminin	V	accord avec espace de séparation
Q	accord alternatif, non binaire	W	accord avec parenthèse
A	mot en anglais	X	accord avec point médian
E	mot épicène	Y	accord avec point
B	mot au verlan	Z	accord avec barre oblique
R	mot raccourci (apocope)	DMF	doublet ordre masculin-féminin
HF	étiquette H/F (homme/femme)	DFM	doublet ordre féminin masculin
S	sans accord genre/nombre		absence d'autres éléments accordables (position 2, 3)
C	féminin marqué en majuscule	0	

Pour ce qui est de la manifestation de l'accord au long d'un seul et même tweet, notre corpus permet de distinguer trois types, à savoir:

- (1) @rMdes_ Pour clore le débat, voici l'excellente analyse de Irène Kaufer. Belle journée à **tu.te.s!** (Belgique)
- (2) @AntTaly vous donne des astuces pour vous équiper comme **un.e vrai.e chercheur.se** en biologie dans la semaine 4 du #MOOCvivant (France)

- (3) Mes **potes** appart d'être **belleau, intelligent/e, drôle**, de me soutenir, de s'aimer mutuellement et de faire partis de ma famille vous servez à quoi ? (je vous aime hein) (Belgique)

Dans les cas comme (1), nous observons un seul élément susceptible d'être inclusif et dont l'inclusion se fait à partir d'un procédé du LI (classifiés comme IS, inclusion simple). Puisqu'il n'y a qu'un seul élément accordé, ce type de tweets n'a pas suffisamment d'information pour analyser la manifestation des accords inclusifs et ils ont été pris en compte seulement pour les analyses globales (figure 1). Les autres analyses ont pris en compte des cas où nous observons deux éléments accordés ou plus et nous les avons divisés en deux groupes: les tweets comme (2) sont classifiés comme des accords uniformes (AU): les éléments accordables présentent un même procédé du LI. Les tweets comme (3) sont classifiés comme accords mixtes (AM): les accords se font avec différents procédés du LI.

3. Résultats

Nos analyses quantitatives prennent en compte la totalité des tweets à inclusion simple (IS), en contraste avec ceux ayant plus d'un élément susceptible d'être accordé, que ce soit des accords uniformes (AU) ou mixtes (AM). Dans le cas spécifique des AM, nous avons aussi observé le positionnement des procédés d'accord LI, spécifiquement dans les tweets des villes de Bruxelles, de Montréal et de Paris, celles dont on avait le plus de données. La figure 1 représente la composition du corpus et la distribution d'accords inclusifs par pays en nombre et en pourcentage :

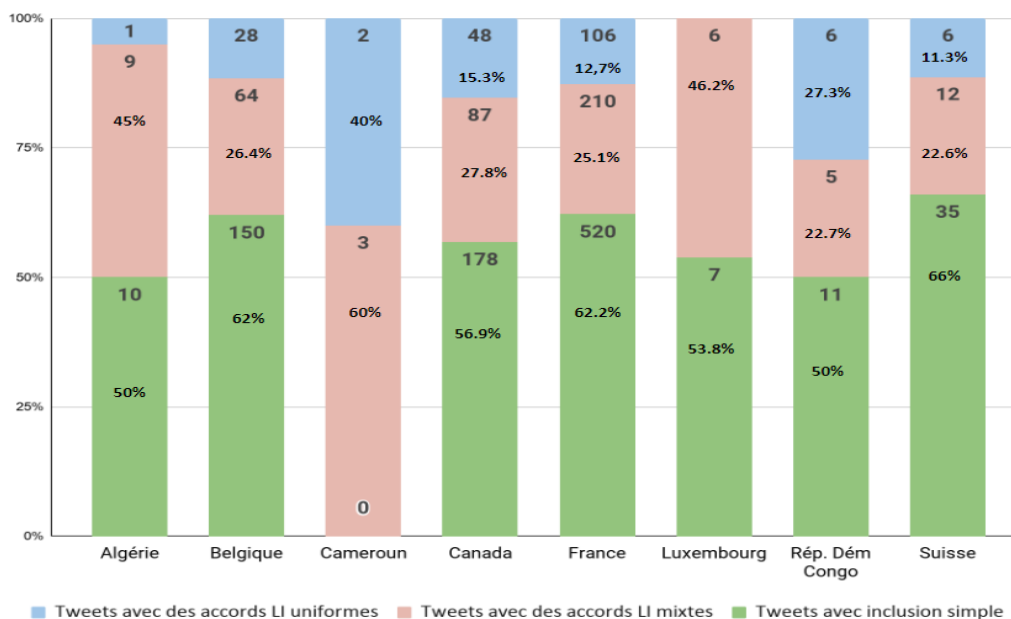


Figure 1. Composition du corpus par pays en nombre et en pourcentage: inclusion simple, accords uniformes et mixtes.

La figure 1 montre que le pourcentage le plus important d'emploi du LI est présent dans les tweets n'ayant qu'un seul élément susceptible d'être inclusif (50% ou plus dans tous les cas sauf le Cameroun). Dans le cas où il y a plus d'un élément susceptible d'être accordé, la tendance est à utiliser des accords mixtes jusqu'à deux fois plus souvent que les accords uniformes. Cette proportion n'est pas observée dans les cas de la République Démocratique du Congo et du Cameroun mais les quantités des données dans ces deux pays sont trop petites pour être considérées dans les analyses.

3.1 Les accords uniformes

Le 33.3% des cas où les tweets présentaient plus d'un élément susceptible d'être accordé ont présenté des accords uniformes, ce qui représente 197/593 occurrences dans tout le corpus. Les exemples (4) à (6) montrent des accords uniformes à partir de l'emploi de parenthèses (4), de barres obliques (5) et de points médians (6):

- (4) Si j'ouvre une école CD il y aura des cours de politique dès la première secondaire, chaque fin d'année l'élève devra proposer et expliquer à toute la classe 5 solutions pour aider la population. **Un(e) bon(ne) président(e)** se forme dès la jeunesse (République Démocratique du Congo).
- (5) Parfois, vous vous comportez comme si on vous avait obligés à follow **un/e tel/le**. Sentez-vous libres d'unfollow, masquer ou bloquer (Cameroun)
- (6) Quand tu as enfin fini #Sorcières de @monachollet, que tu vas le prêter à **tou•tes tes ami•es**, lancer une pétition pour qu'il soit obligatoirement enseigné à l'école, et porter chapeau pointu et cape à toutes les prochaines manif #féministes (France)

La figure 2 représente les occurrences d'AU par pays :

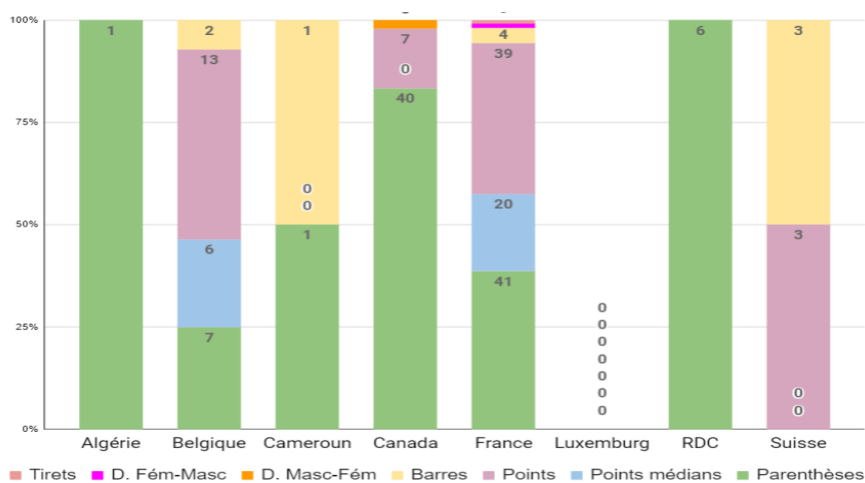


Figure 2. Distribution d'accords uniformes par pays.

Pour les AU, la tendance générale dans notre corpus est l'emploi de parenthèses. Pour le Canada, c'est le procédé dominant pour les AU, avec 40 sur 48 occurrences. Dans le cas de France, les 106 occurrences d'AU se divisent de manière presque égale entre 41 cas d'emploi de parenthèses et 39 d'emploi de points, plaçant le point médian dans la troisième position d'emploi avec 20 occurrences. Pour ce qui est de la Belgique, les 28 AU se divisent entre 13 emplois de points, 7 d'emploi de parenthèses et 6 des points médians. Les données pour les autres pays sont trop éparpillées pour être prises en compte.

3.2 Les accords mixtes

Le 66.7% de cas où les textes avaient plus d'un élément susceptible d'être accordé ont présenté des accords mixtes, ce qui représente 396/593 occurrences dans tout le corpus (voir figure 1). Étant donné la complexité de cette mixité et les limites de cette étude exploratoire, nous avons abordé uniquement deux axes de cette mixité : la diversité de combinaisons et le positionnement de ces procédés dans la distribution du tweet.

Pour ce qui est la diversité de procédés utilisés dans les AM, observons quelques exemples de combinaison trouvées dans notre corpus. Dans (7), il y a une combinaison de forme abrégée avec point (*un.e*), suivi d'une forme épiciène (*pote*) et un pronom d'objet direct doublé en ordre féminin-masculin (*lale*) :

- (7) @IamMuruka Je suis d'accord. Mais ce n'était pas le moment. [...]. Une vexation perso ça vaut pas le coup de prendre le risque de foutre en l'air un projet de vie. Moi **un.e pote** qui me fait ça c fini terminé, je **lale** raye de ma vie. (Suisse).

Dans le cas de (8), nous pouvons observer l'emploi d'une néoforme pour un pronom personnel sujet de troisième personne singulier (*iel*), suivi d'un participe passé accordé au masculin singulier, accompagné d'un adjectif en anglais comme forme épiciène :

- (8) @mrwnb ça vnr genre **iel** est **né woke**???? ça arrive de dire des choses maladroites surtout quand t'as 15ans et que t'y connais presque rien mdr (France)

Dans le cas de (9), nous observons un doublet complet en ordre masculin-féminin (*il ou elle*), suivi d'un adjectif accordé au féminin (*jalouse*), ce qui pourrait être l'accord de proximité (mais aussi de majorité ou d'énonciation, un exemple des raisons pour lesquelles nous avons décidé de ne pas utiliser ces classifications dans cette étude).

- (9) @ineselg_ Au moins ya dit stp et merci serieux **il ou elle** est juste **jalouse**, tes apparue dans mon feed, je te suis pas, mais tu est belle et tu a juste bruisé son ego par ta beauté (Canada)

D'autres occurrences d'accord mixte observées présentent aussi d'autres pronoms alternatifs (10), l'emploi de l'étiquette (F/H) ou (H/M) après un métier pour marquer l'inclusivité (11), entre d'autres:

- (10) @hellnoctopus Je te soutiens, et je soutiens **tout.e.s ceux concerné.e.s** pas ses propos horribles Vous êtes fort.e.s (France)
- (11) @AsteropFr. Asterop recherche Chef de projet (**F/H**). Ce poste à temps plein (CDI) est basé à Ivry-sur-Seine (France).

La taille de notre corpus étant trop réduite pour donner des conclusions significatives concernant ces combinaisons, nous nous permettons uniquement de présenter quelques tendances observées dans les trois pays où nous avons obtenu les plus de données: la Belgique (Bruxelles), le Canada (Montréal) et la France (Paris).

3.2.1 Belgique

Nous avons classifié 64 accords mixtes dans l'ensemble de 92 tweets belges dans notre corpus :

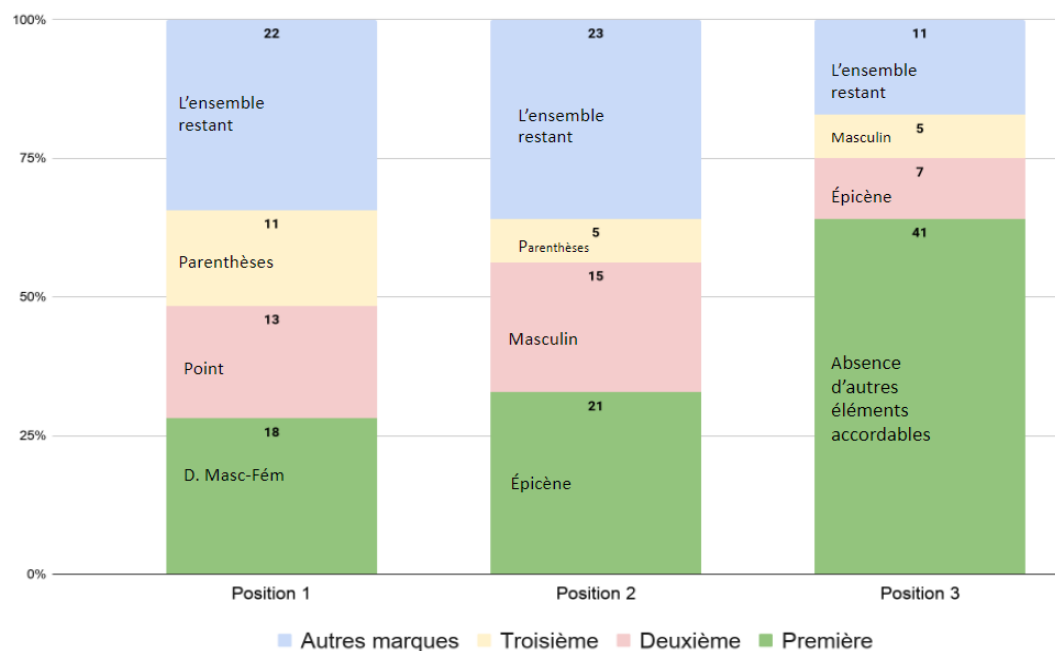


Figure 3. Accords mixtes dans les tweets belges: premier, deuxième et troisième procédé LI les plus utilisés dans les trois premières positions.

Cet ensemble présente, dans le premier élément accordable, un accord au moyen de doublets en ordre masculin-féminin (18 occurrences), suivi par l'emploi des points et de parenthèses, avec 13 et 11 occurrences respectivement. Pour ce qui est du deuxième élément accordable, l'emploi des formes épiciennes est majoritaire (21), suivi par l'accord au masculin (15) et l'accord inclusif avec des parenthèses (5).

3.2.2 Canada

Nous avons classifié 87 accords mixtes dans l'ensemble de 135 tweets canadiens dans notre corpus. Le premier élément accordable dans les tweets de Montréal dans notre corpus présente un accord au moyen de parenthèses (31 occurrences), suivi par l'emploi de doublets en ordre masculin-féminin (19) et de points (10). Pour ce qui est du deuxième élément accordable, l'emploi des formes épiciènes a été équivalent à l'emploi de doublets en ordre masculin-féminin, chaque procédé avec 21 occurrences dans le corpus canadien, suivis par l'accord au masculin (15).

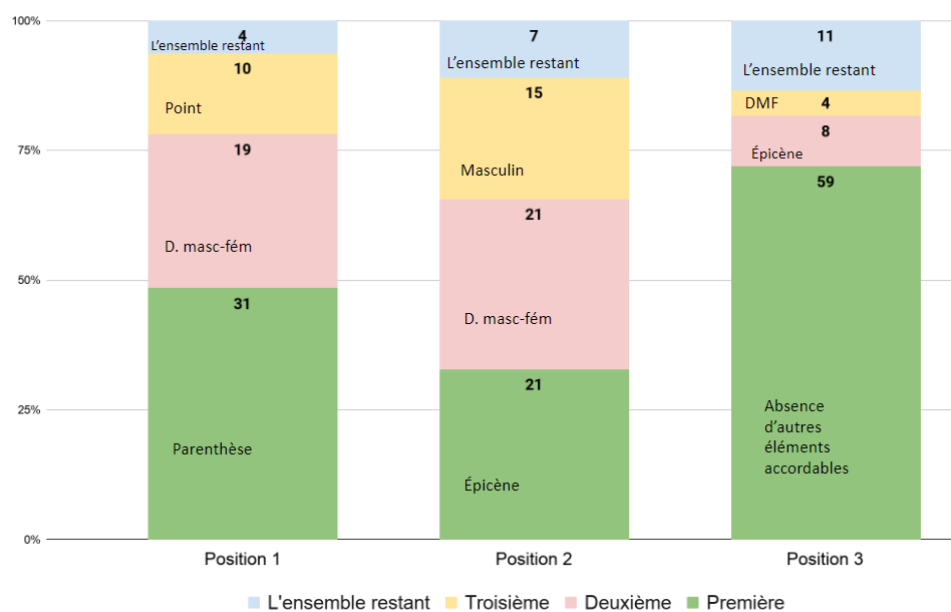


Figure 4. Accords mixtes dans les tweets canadiens: premier, deuxième et troisième procédé LI les plus utilisés dans les trois premières positions.

L'ensemble des autres procédés constitue un chiffre bas, avec quatre pour la première position et sept dans la deuxième, ce qui pourrait indiquer une certaine régularité dans l'emploi de formes mentionnées dans ces positions.

3.2.3 France

Nous avons classifié 210 accords mixtes dans l'ensemble de 316 tweets français dans notre corpus. Pour le cas de Paris, nous avons observé les cinq premiers procédés utilisés dans les trois premiers éléments accordables. Le premier élément accordable dans les tweets de Paris dans notre corpus présente un accord au moyen des doublets en ordre masculin-

féminin (58 occurrences), suivi par l'emploi des parenthèses (36), de points (32), de doublets féminin-masculin (26) et de néoformes (16). Pour ce qui est du deuxième élément accordable, l'emploi des formes épïcènes est en tête (58), suivi par l'accord au masculin (41) et l'accord inclusif avec des parenthèses (17), des accords au féminin (15) et de néoformes (13).

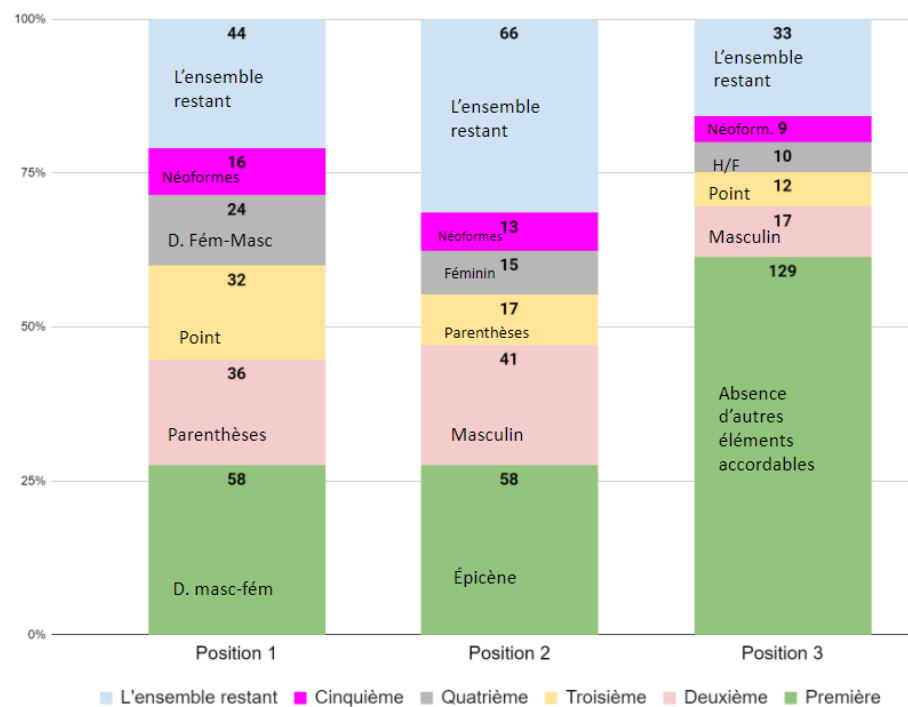


Figure 5. Accords mixtes dans les tweets français: premier, deuxième et troisième procédé LI les plus utilisés dans les cinq premières positions.

Il est important de souligner que dans le cas de Belgique et de France l'ensemble des autres procédés constitue le chiffre d'emploi le plus élevé dans les deux premières positions: 22 occurrences pour la Belgique et 44 pour la France dans la première position; 23 pour la Belgique, 66 pour la France pour la deuxième position, ce qui pourrait indiquer qu'il n'y a pas de stabilisation dans le choix de combinaisons mais une situation de mixité complexe et dynamique. Un nombre important de tweets n'a pas de troisième élément accordable: 41 occurrences dans le cas de Belgique, 59 dans le cas de Canada, 129 dans le cas de France. De ceux qui en ont un troisième élément: le procédé inclusif le plus utilisé est la forme épïcène en Belgique (sept) et au Canada (huit) et l'accord au masculin dans le cas de France (17).

4. Discussion des résultats

Netlytic fournit des données en grande quantité et de manière automatisée, ce qui est un avantage pour ce type de recherches. Cependant il faut continuer à perfectionner des mécanismes automatisés ou semi-automatisés pour le lourd travail de dépouillement pour les doublons, les tweets à localisation absente, bloquée ou même aberrante, comme le cas de tweets de Kinshasa recueillis avec des paramètres de localisation pour Hanoï. On a dû aussi laisser de côté la dimension diachronique dans notre étude étant donné les restrictions du logiciel. Les résultats de cette expérience nous ont permis aussi d'observer le comportement des descripteurs pour recueillir un corpus de ce type à d'avoir une approximation à ceux qui seraient les plus productifs dans des expériences ultérieures. Dans ce sens, les groupes de descripteurs les plus significatifs ont été ceux qui contenaient des formes abrégées: les pronoms, les adjectifs et les articles indéfinis (276 tweets, groupe 1, tableau 1) et les substantifs reliés à des métiers (193 tweets, groupe 7, tableau 1).

Même si on visait les emplois inclusifs non-institutionnels, nous observons qu'il y a un ensemble de tweets provenant des comptes (sémi-)institutionnels. Nous trouvons par exemple dans notre corpus un ensemble d'annonces d'emploi issus d'entités capables d'employer des gens. On observe toutefois que même s'il peut s'agir d'institutions, ces textes semblent moins soignés et plus informels que le langage qui serait utilisé dans d'autres plateformes, voire dans l'imprimé.

Dans notre corpus, les doublets en ordre masculin-féminin sont le procédé d'accord inclusif le plus utilisé dans le premier élément accordable des tweets provenant de Bruxelles et de Paris, tandis qu'à Montréal la préférence dans cette position est l'accord par parenthèses. Les formes épïcènes sont le procédé inclusif le plus utilisé pour le deuxième élément accordable dans les trois villes en question. Cependant, on observe qu'à Montréal cet emploi est équivalent en quantité à l'emploi de doublets en ordre masculin-féminin, et qu'à Paris et à Bruxelles le chiffre total d'autres procédés est élevé dans cette position, avec 66 et 23 cas respectivement. En général, les accords mixtes ne présentent pas de troisième élément accordable dans notre corpus, ce que nous observons à partir de 41 occurrences à Bruxelles, 59 à Montréal et 129 à Paris.

De manière préliminaire, nous observons que l'élément en première position est marqué soit par des doublets complets ou abrégés. Dans le cas des abrégés, les parenthèses (au Canada) et les points (au Belgique et en France) sont les procédés typographiques préférés. Pour ce qui est de l'élément en deuxième position, nous remarquons le passage vers des formes inclusives moins marquées, comme les mots épïcènes, mais aussi les doublets complets. Cependant il faut aussi souligner un pourcentage important de formes accordées au masculin à Bruxelles et à Paris. L'absence d'un troisième élément à accorder

pourrait être due à une volonté d'éviter les marques d'accord mais aussi à la brièveté caractéristique des tweets. Cette question sera reprise dans nos études ultérieures.

5. Conclusions préliminaires

Il semble y avoir un ensemble de systèmes optionnels de concordance qui se chevauchent, chacun d'entre eux comme une alternative à l'utilisation traditionnelle du masculin dit 'générique'. Nous observons aussi une diversité régionale dans les différents territoires francophones, ce qui ne peut pas être considéré comme une variation stable mais qui ressemble plus à une étape de changement ou d'ajustement (hypothèse à vérifier). Le fait d'avoir observé un pourcentage entre 22% et 60% d'accords mixtes dans notre corpus semble nous indiquer qu'il ne s'agit pas de simples "erreurs" mais bien de choix des locutrices et locuteurs. Il semble donc que les accords mixtes seraient une stratégie consciente, raison pour laquelle nous proposons cette mixité¹ comme un des traits caractéristiques du LI dans l'actualité (Diaz, en préparation).

Pour ce qui est du but des accords mixtes dans le contexte du LI, nous considérons que les arguments de Kalinowski (2019) à propos du LI en espagnol argentin peuvent aussi englober et justifier ce trait de mixité : « le LI est un phénomène rhétorique-discursif plutôt que grammatical dont l'objectif ne vise pas le changement linguistique, plutôt le changement de certaines réalités sociales » (2019: 53). On ne devrait donc s'attendre à un usage 'uniforme', grammaticalisé ni généralisé puisque ce ne serait pas le but. Ceci ne serait pas non plus un « péril mortel » pour la langue, comme le caractérise l'Académie française (2017). C'est un phénomène, pour l'instant, minoritaire dans le français comme ailleurs, mené par des mouvements et des acteurs sociaux ayant accès à des réseaux sociaux tels que Twitter.

Dans ce cadre, nous considérons, comme Alpheratz, que le langage inclusif peut être assumé comme une variation diaéthique, plutôt que diatopique, diastratique ou diaphasique, car « elle relève d'autre chose, d'un acte politique motivée par une conscience de genre et par le souci du respect des représentations symboliques et des catégories sociales minorisées » (2018: 7). Le LI fait ainsi appel à une population qui veut se montrer rebelle vis-à-vis de l'école, aux institutions, aux académies, et qui, pour le faire, profite d'une dimension disons ludique de la langue, là où les espaces tels que Twitter le

¹ Au-delà de notre approche qui vise le contexte non-institutionnel, la mixité des accords inclusifs est aussi présente ailleurs. Même le congrès annuel de l'Association canadienne de linguistique 2020 emploie dans ses instructions (<http://cla-acl.ca/congres-de-2020-meeting/>) un ensemble mixte de procédés : des formes épïcènes, des doublets complets et aussi des abrégés au moyen de barres obliques, de traits d'union, etc. La coexistence de différents procédés inclusifs montrerait un certain niveau d'acceptation institutionnelle.

permettent. La mixité des accords inclusifs pourrait, dans ce sens, représenter une voie opposée à ce qui est normatif et une recherche de représentation de la richesse de la diversité.

Références

- Abbou, Julie. 2017. (Typo)graphies anarchistes. Où le genre révèle l'espace politique de la langue. *Les langages du politique*, (1):53-72.
- Académie française. 2017. *Déclaration de l'académie française sur l'écriture dite «inclusive» adoptée à l'unanimité de ses membres dans la séance du jeudi 26 octobre 2017*. <http://www.academie-francaise.fr/actualites/declaration-de-lacademie-francaise-sur-lecriture-dite-inclusive>
- Alpheratz. 2018. Français inclusif : conceptualisation et analyse linguistique. *SHS Web of Conferences*, vol. 46, EDP Sciences. Jan. 2018, 13003-, doi:10.1051/shsconf/20184613003.
- Alpheratz. *Grammaire du français inclusif*. Villedieu-sur-Indre : Vent solars.
- Bouzouita, Miriam et Enrique Pato. 2019. Por qué no gustas de mí como yo gusto de ti? El verbo 'gustar' y su complemento preposicional en español actual. *Circulo de lingüística aplicada a la comunicación*, 79:161–86.
- Brauer, Markus, et Michaël Landry. 2008. Un ministre peut-il tomber enceinte? l'impact du générique masculin sur les représentations mentales. *L'Année psychologique*, 108(2):243-272
- Burnett, Heather et Céline Pozniak, 2020. Political Dimensions of Écriture Inclusive in Parisian Universities. Ms. Université de Paris, LLF, CNRS.
- Diaz, Yarubi. (à paraître). « Langage inclusif et accord variable en français et espagnol sur Twitter ». Thèse de doctorat, Université Western Ontario.
- Dister, Anne et Marie-Louise Moreau. 2009. Histoire de la féminisation. *Féminiser? vraiment pas sorcier: La féminisation des noms de métiers, fonctions, grades et titres*. 9-23. Bruxelles: De Boeck/Duculot.
- Escalante, Chelsea. 2015. ¿Qué twiteastes tú? Variation in second person singular preterit –s in Spanish tweets. *Entrehojas: Revista de Estudios Hispánicos* 5(1):1-18.
- Haddad, Raphaël et Carline Baric. 2016. Manuel d'écriture inclusive. Faites progresser l'égalité des femmes.hommes par votre manière d'écrire. *Mots-Clés*. <http://www.ecriture-inclusive.fr/>
- InternetWorldStats, 2020. *World Internet Usage and Populations Statistics 2020*. <https://www.internetworldstats.com/stats.htm>
- Kalinowski, Santiago. 2019. Lenguaje inclusivo: cambio lingüístico o cambio social. *Revista del Colegio de Traductores Públicos de la Ciudad de Buenos Aires*. 141:53-55.
- La vie en queer, 2018. Petit dico de français neutre/inclusif. *La Vie en Queer* (blogue). 26 juillet 2018. <https://lavieenqueer.wordpress.com/2018/07/26/petit-dico-de-francais-neutre-inclusif/>
- Marttinen Larsson, Matti et Miriam Bouzouita. 2018. Encima de mí vs. encima mío: Un análisis variacionista de las construcciones adverbiales locativas con complementos preposicionales y posesivos en Twitter. *Moderna sprak*. 112(1): 1–39.
- Observatoire de la francophonie. 2020. *La langue française dans le monde, Édition de 2019*. Gallimard, http://observatoire.francophonie.org/wp-content/uploads/2020/02/Edition-2019-La-langue-francaise-dans-lemonde_VF-2020-.pdf
- Slemp, Katie, David Heap et Yarubi Diaz (à paraître). Writing vs. Speech: How do Spanish online videos pronounce gender inclusive language? Dans *Actes du congrès annuel de l'Association canadienne de linguistique 2019*, ed. Katie Martin, 14 pp. Toronto : Association canadienne de linguistique.
- Viennot, Eliane, Maria Candea, Yannick Chevalier, Sylvia Duverger et Anne Marie Houdebine. 2015. Les offenses. Dans *L'Académie contre la langue française : le dossier féminisation*, 65-78. Editions iXe, Donnemarie-Dontilly.